Rate of Convergence

The sequences \( w_n = 1/\log_2(n) \), \( x_n = 1/n \), \( y_n = 1/n^2 \), and \( z_n = 1/2^n \) all converge to 0. How fast? Well, \( w_8 = 1/3 \), \( x_8 = 1/8 \), \( y_8 = 1/64 \), and \( z_8 = 1/256 \) is true, but doesn’t really convey how slow \( w_n \to 0 \), or how fast \( z_n \to 0 \). We use functions like these, and more generally \( 1/n^p \) and \( 1/e^n \) as yardsticks (or benchmarks) with which to compare the speed of convergence of algorithms.

In our context, we are usually trying to compute better and better approximations \( p_n \) to some value \( p \), and we want to know how fast the error \( e_n = |p - p_n| \) is converging toward 0. It is nice to be able to say \( p_n \to p \) (our approximations will eventually be close to the answer), but will it take a second or a week? In the example above \( w_{512} > 10^{-1} \), but \( z_{30} < 10^{-9} \).

We will use these standard sequences above as benchmarks. If a sequence converges about as slow as the \( w_n \), we will say it has a logarithmic rate of convergence, but if it goes fast like \( z_n \), we will say its rate of convergence is exponential.

Here is the formal definition. Think of \( \beta_n \) as one of the sequences given above. If \( \{\beta_n\} \) is a positive sequence converging to 0, then we say that \( p_n \to p \) with rate of convergence \( \beta_n \) if \( \exists N \) and \( k > 0 \) such that \( \forall n > N, \)

\[
|p_n - p| \leq k\beta_n.
\]

In this case we write \( p = p_n + O(\beta_n) \).

Now suppose \( p_7 \approx p \). What does that mean? What about \( p_{18} \)? Well, if we can write \( p_n = p + O(1/n^3) \), then we know that \( p_n \to p \) at least as fast as \( k/n^3 \to 0 \), for some constant \( k \), and it is probably safe to say that \( p_{18} \) is about \((18/7)^3 \approx 17 \) times more accurate than is \( p_7 \).

In the same way that we have just measured the error in a discrete setting, we can measure the error associated with a continuous parameter. Suppose \( \lim_{x \to 0} f(x) = L \). How fast? We say that \( f(h) \to L \) as \( h \to 0 \) with rate of convergence \( h^q \) if \( \exists \delta \) and \( k > 0 \) such that \( \forall |h| \leq \delta, \)

\[
|f(h) - L| \leq kh^q.
\]

We write \( f(h) = L + O(h^q) \). The idea here is that we are approximating \( L \) by \( f(h) \), and we would like to compare \( |f(h) - L| \) to \( h^q \), because we have a feeling for how fast \( h^q \to 0 \).

For example, \( \lim_{x \to 0} \sin(x) = 0 \). So for small \( x \), \( \sin(x) \approx 0 \). But we can do better:

\[
\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots,
\]

so we can write \( \sin(x) = 0 + O(x) \), or to give more information \( \sin(x) = x + O(x^3) \), or even more: \( \sin(x) = x - x^3/6 + O(x^5) \), etc.